MMSP 2018 Technical Program

Vancouver, Canada, August 29-31, 2018

Conference venue
SFU Segal Building
500 Granville Street
Vancouver, BC, V6C 1W6
Canada

	Wednesday August 29, 2018	Thursday August 30, 2018	Friday August 31, 2018
8:45am – 9:00am	Opening		
9:00am – 10:00am	Keynote 1	Keynote 2	Keynote 3
10:00am – 10:20am	Break with refreshments	Break with refreshments	Break with refreshments
10:20am – 12:10pm	Special Session SS.W1	Special Session SS.T1	Special Session SS.F1
12:10pm – 1:30pm	Lunch	Lunch MMSP TC meeting	Lunch
1:30pm – 2:30pm	Poster Session PS.W1	Poster Session PS.T1	Poster/Demo Session PS.F1
2:30pm – 2:50pm	Break with refreshments	Break with refreshments	Break with refreshments
2:50pm – 4:20pm	Lecture Session LS.W1	Lecture Session LS.T1	Lecture Session LS.F1
4:20pm – 4:30pm	Break with refreshments	Break with refreshments	Closing
4:30pm – 6:00pm	Lecture Session LS.W2	Lecture Session LS.T2	

Wednesday, August 29, 2018

Keynote 1
Time: 9:00am – 10:00am

Title: **Giving Creatives Bigger Lever**

Gordon Durity
Executive Audio Director, Electronic Arts

Bio: Gordon Durity has over 30 years experience in creating soundtracks, writing songs and producing audio in the areas of film, television, games, multimedia, and album production. He currently serves as an Executive Audio Director at Electronic Arts, one of the largest developers of video games in the world. There Gordon designs and supervises the creation and implementation of audio for various titles, works with technical staff to develop new cutting edge technologies, sits on game design groups, and supervises audio at various EA Studios locations worldwide. Gordon currently serves as a consultant to educational institutions, film and music industries on how best to maximize opportunities in the exploding games and multimedia industries. He has also been a guest lecturer at Simon Fraser University, University of British Columbia and Emily Carr University of the Arts. As a composer, Gordon has written scores for feature and short films, television, radio, dance, theatre, runway shows and games. Gordon has produced and written with many international artists and has had songs charted all over the world. He has worked for USA Network, Fox, Disney, Paramount, CBC, BBC, Trimark, NFB, PBS, ABC, and many independents. Gordon's current research focus is on audio and experience design and implementation for virtual environments for Virtual Reality/Augmented Reality, and machine learning-based speech synthesis.

Abstract: We are in a place in the interactive entertainment industries where content is king. Games and apps are no longer single delivered products, but are becoming more like an OS by which an ongoing stream of content can be delivered to the consumer. We are also dealing with multiple platforms from mobile devices to traditional consoles to VR/AR rigs. As content producers we are researching, applying and integrating machine learning, AI, and procedural/generative systems into our development processes to amplify the capabilities of the modern day "cyber-artist." This presentation will explore and examine some of the thinking and work we are doing at Electronic Arts in these areas, and specifically in the realm of audio.

Special Session SS.W1: Spatial Acoustics Using Sensor Arrays: Visions to Implementations
Session Chair: Rodney G. Vaughan (Simon Fraser University)
Time: 10:20am – 12:10pm

ID Title and Authors

30 *A Planar Microphone Array for Spatial Coherence-Based Source Separation*
Abdullah Fahim, Prasanga Samarasinghe, Thushara Abhayapala, and Hanchi Chen (Australian National University)

34 *Sound Environment Reproduction for Health and Safety Studies Using Microphone Arrays, Wave Field Synthesis and the Lasso Minimizer*
Philippe-Aubert Gauthier and Alain Berry (Université de Sherbrooke)

75 *Speech Intelligibility of Microphone Arrays in Reverberant Environments with Interference*
Elham Ideli, Rodney G. Vaughan, and Ivan V. Bajić (Simon Fraser University)

87 *Beamforming with Partial Knowledge of the Acoustic Scenario*
W. Bastiaan Kleijn (Victoria University of Wellington), Christopher Laguna, Alejandro Luebs, Andrew MacDonald, and Jan Skoglund (Google)

92 *Multiple Source Location Estimation on a Dataset of Real Recordings in a Wireless Acoustic Sensor Network*
Anastasios Alexandridis (University of Crete), Anthony Griffin (Auckland University of Technology), and Athanasios Mouchtaris (University of Crete)

Poster Session PS.W1: Multimedia Processing and Analysis I
Session Chair: Eduardo Peixoto (Universidade de Brasília)
Time: 1:30pm – 2:30pm

ID Title and Authors

22 *Online Music Performance Tracking Using Parallel Dynamic Time Warping*
I-Chieh Wei and Li Su (Institute of Information Science, Academia Sinica)

26 *A Hybrid DSP/Deep Learning Approach to Real-Time Full-Band Speech Enhancement*
Jean-Marc Valin (Mozilla Corporation)

71 *A Sub-Aperture Image Selection Refinement Method for Progressive Light Field Transmission*
Wallace de Souza, Bruno Macchiavello, Edson Hung, Eduardo Peixoto (University of Brasilia), and Gene Cheung (National Institute of Informatics, Japan)

81 *Robust Polyphonic Sound Event Detection by Using Multi Frame Size Denoising Autoencoder*

Jianchao Zhou, Xiaoou Chen, and Deshun Yang (Peking University)

82 *Blastomere Cell Counting and Centroid Localization in Microscopic Images of Human Embryo*
Reza Moradi Rad, Parvaneh Saeedi (Simon Fraser University), Jason Au, and Jon Havelock (Pacific Centre for Reproductive Medicine)

98 *Spatial Reinforcement and Immersive Audio*
Timothy Bartoo (Harmonic Functions), Robin Whittaker, and Dave Haydon (Outboard Electronics)

142 *Reliability Analysis of IoVT Based Intelligent Video Surveillance System*
Tanin Sultana, Mohammad Wajih Alam, and Khan A. Wahid (University of Saskatchewan)

156 *Deep Transfer Learning for Hyperspectral Image Classification*
Jianzhe Peter Lin, Z. Jane Wang, and Rabab Ward (University of British Columbia)

158 *MV-YOLO: Motion Vector-Aided Tracking by Semantic Object Detection*
Saeed Ranjbar Alvar and Ivan V. Bajić (Simon Fraser University)

Lecture Session LS.W1: Visual Indexing, Analysis, and Representation
Session Chair: Parvaneh Saeedi (Simon Fraser University)
Time: 2:50pm – 4:20pm

ID Title and Authors

7 *Feature Fusion for Robust Patch Matching with Compact Binary Descriptors*
Andrea Migliorati (Università degli Studi di Brescia), Attilio Fiandrotti (Politecnico di Torino), Gianluca Francini (Telecom Italia S.P.A), Skjalg Lepsøy (Telecom Italia S.P.A), and Riccardo Leonardi (Università degli Studi di Brescia)

33 *Efficient Object Tracking in Compressed Video Streams with Graph Cuts*
Fernando Bombardelli (Fraunhofer HHI), Serhan Gül (Fraunhofer HHI), Daniel Becker (Daimler Center for Automotive IT Innovations), Matthias Schmidt (Daimler Center for Automotive IT Innovations), and Cornelius Hellge (Fraunhofer HHI)

76 *Sparse Hartley Modeling for Fast Image Extrapolation*
Nils Genser, Simon Grosche, Jürgen Seiler, and Andre Kaup (Friedrich-Alexander-Universität Erlangen-Nürnberg)

94 *CPNet: A Context Preserver Convolutional Neural Network for Detecting Shadows in Single RGB Images*
Sorour Mohajerani and Parvaneh Saeedi (Simon Fraser University)

159 *A Cloud Detection Algorithm for Remote Sensing Images Using Fully Convolutional Neural Networks*
Sorour Mohajerani, Parvaneh Saeedi, and Thomas A. Krammer (Simon Fraser University)

Lecture Session LS.W2: Multimedia Compression
Session Chair: Ivan V. Bajić (Simon Fraser University)
Time: 4:30pm – 6:00pm

ID Title and Authors

37 *Test Zonal Search based on Region Label (TZSR) for Motion Estimation in HEVC*
Iris C. Linck (University of Colorado, Denver), Arthur Gomez (CNPq, Brasilia), and Gita Alaghband (University of Colorado, Denver)

46 *Motion Compensated Prediction for Translational Camera Motion in Spherical Video Coding*
Bharath Vishwanath, Kenneth Rose, and Tejaswi Nanjundaswamy (University of California, Santa Barbara)

93 *Deep Network Based Image Compression with Adaptive Pre- and Postprocessing*
Shurun Wang, Zhenghui Zhao, Chuanmin Jia, Xiang Zhang (Peking University), Xinfeng Zhang (University of Southern California), Shanshe Wang, Siwei Ma, and Wen Gao (Peking University)

151 *Near-Lossless Deep Feature Compression for Collaborative Intelligence*
Hyomin Choi and Ivan V. Bajić (Simon Fraser University)

157 *Downsampling Based Image Coding Using Dual Dictionary Learning and Sparse Representations*
Ali Akbari and Maria Trocan (Institut Supérieur d'Electronique de Paris)

Thursday, August 30, 2018

Keynote 2
Time: 9:00am – 10:00am

Title: **Coordinated Dynamic Mining of 3D Physical World from Videos**

Jenq-Neng Hwang
Professor, EE Department, University of Washington

Bio: Dr. Jenq-Neng Hwang received the BS and MS degrees, both in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1981 and 1983 separately. He then received his Ph.D. degree from the University of Southern California. In the summer of 1989, Dr. Hwang joined the Department of Electrical Engineering of the University of Washington in Seattle, where he has been promoted to Full Professor since 1999. He served as the Associate Chair for Research from 2003 to 2005, and from 2011-2015. He is currently the Associate Chair for Global Affairs and International Development in the EE Department. He has written more than 330 journal, conference papers and book chapters in the areas of multimedia signal processing, and multimedia system integration and networking, including an authored textbook on 'Multimedia Networking: from Theory to Practice', published by Cambridge University Press. Dr. Hwang has close working relationship with the industry on multimedia signal processing and multimedia networking. Dr. Hwang received the 1995 IEEE Signal Processing Society's Best Journal Paper Award. He is a founding member of Multimedia Signal Processing Technical Committee of IEEE Signal Processing Society and was the Society's representative to IEEE Neural Network Council from 1996 to 2000. He is currently a member of Multimedia Technical Committee (MMTC) of IEEE Communication Society and also a member of Multimedia Signal Processing Technical Committee (MMSP TC) of IEEE Signal Processing Society. He served as associate editors for IEEE T-SP, T-NN and T-CSVT, T-IP and Signal Processing Magazine (SPM). He is currently on the editorial board of ZTE Communications, ETRI, IJDMB and JSPS journals. He served as the Program Co-Chair of IEEE ICME 2016 and was the Program Co-Chairs of ICASSP 1998 and ISCAS 2009. Dr. Hwang is a Fellow of IEEE since 2001.

Abstract: With the huge amount of networked video cameras available everywhere nowadays, such as the statically deployed surveillance cameras or the constantly moving cameras on the vehicles or drones, there is an urgent need of systematic and coordinated mining of the dynamic environment in the 3D physical world, so that the explored information can be exploited for various smart city applications, such as security surveillance, intelligent transportation, business statistics collection, health monitoring of communities, and etc. In this talk, I will first present an automated and robust human/vehicle tracking directly in 3D space through self-calibration of static and moving monocular cameras. These tracked objects locations and speed, as well as their poses, can all be described based on the GPS coordinates, so that the tracked objects from multiple cameras can then be effectively integrated and

reconstructed in the 3D real-world space for many smart city and intelligent transportation applications.

Special Session SS.T1: Recent Advances in Image Restoration and Quality Metrics for Restoration Algorithms
Session Chairs: Purang Abolmaesumi (University of British Columbia)
Time: 10:20am – 12:10pm

ID Title and Authors

38 *Enhanced Steerable Pyramid Transformation for Medical Ultrasound Image Despeckling*
Prerna Singh, Ramakrishnan Mukundan (University of Canterbury), and Rex De Ryke (Canterbury District Health Board)

41 *Color-Guided Depth Map Super-Resolution via Joint Graph Laplacian and Gradient Consistency Regularization*
Rong Chen, Deming Zhai, Xianming Liu, and Debin Zhao (Harbin Institute of Technology)

89 *Reduced-Reference Image Quality Assessment Based on Free-Energy Principle with Multi-Channel Decomposition*
Wenhan Zhu, Guangtao Zhai, and Xiaokang Yang (Shanghai Jiao Tong University)

102 *Convolutional Neural Network Based Intermediate View Synthesis for Light Field Image Compression*
Yekang Yang, Zhenghui Zhao, Chuanmin Jia, Xiang Zhang, Shanshe Wang, and Siwei Ma (Peking University)

112 *A Large-Scale Compressed 360-Degree Spherical Image database: From Subjective Quality Evaluation to Objective Model Comparison*
Wei Sun (Shanghai Jiao Tong University), Ke Gu (Beijing University of Technology), Siwei Ma (Peking University), and Guangtao Zhai (Shanghai Jiao Tong University)

139 *ConvCSNet: A Convolutional Compressive Sensing Framework Based on Deep Learning*
Xiaotong Lu, Weisheng Dong, Peiyao Wang, Guangming Shi, and Xuemei Xie (Xidian University)

Poster Session PS.T1: Multimedia Processing and Analysis II
Session Chair: George Tzanetakis (University of Victoria)
Time: 1:30pm – 2:30pm

ID Title and Authors

39 *Video Classification of Farming Activities with Motion-Adaptive Feature Sampling*
He Liu, Amy R. Reibman, Aaron C. Ault, and James V. Krogmeier (Purdue University)

53 *A New Retrieval System Based on Low Dynamic Range Expansion and SIFT Descriptor*
Raoua Khwildi and Azza Ouled Zaid (Université de Tunis)

61 *Non-local Super Resolution in Ultrasound Imaging*
Parviz Khavari, Amir Asif, and Hassan Rivaz (Concordia University)

78 *Image Forensics in Online News*
Federica Lago, Quoc-Tin Phan, and Giulia Boato (University of Trento)

90 *A Dual Path Deep Network for Single Image Super-Resolution Reconstruction*
Fateme Mirshahi and Parvaneh Saeedi (Simon Fraser University)

95 *SPmat: A Framework and Data Representation for Binary Image Processing*
Fabrizio Pedersoli and George Tzanetakis (University of Victoria)

99 *Identifying Image Provenance: An Analysis of Mobile Instant Messaging Apps*
Quoc-Tin Phan (University of Trento), Cecilia Pasquini (Universität Innsbruck), Giulia Boato, and Francesco De Natale (University of Trento)

130 *Quality Assessment of Deep-Learning-Based Image Compression*
Giuseppe Valenzise (Université Paris-Sud), Andrei Purica (Telecom ParisTech), Vedad Hulusic (Bournemouth University), and Marco Cagnazzo (Telecom ParisTech)

143 *Fast 3D Point Cloud Denoising via Bipartite Graph Approximation & Total Variation*
Chinthaka Dinesh (Simon Fraser University), Gene Cheung (National Institute of Informatics, Japan), Ivan V. Bajić (Simon Fraser University), and Cheng Yang (National Institute of Informatics, Japan)

Lecture Session LS.T1: Multimedia Processing, Forensics, and Analysis
Session Chair: Z. Jane Wang (University of British Columbia)
Time: 2:50pm – 4:20pm

ID Title and Authors

27 *Image Inpainting Detection Based on a Modified Formulation of Canonical Correlation Analysis*
Xiao Jin (Tianjin University), Yu-ting Su (Tianjin University), Yongwei Wang (University of British Columbia), and Z. Jane Wang (University of British Columbia)

60 *Color Noise-Based Feature for Splicing Detection and Localization*
Christophe Destruel, Vincent Itier, Olivier Strauss, and William J.-P. Puech (Université de Montpellier)

84 *An Adaptive Bandpass Filter based on Temporal Spectrogram Analysis for Photoplethysmography Imaging*
Timon Blöcher, Kai Zhou, Simon Krause (FZI Forschungszentrum Informatik), and Wilhelm Stork (Kalrsruhe Institute of Technology)

121 *Decoding Music in the Human Brain using EEG Data*
Chris Foster, Dhanush Dharmaretnam, Haoyan Xu, and George Tzanetakis (University of Victoria)

133 *Improving Real-time Pedestrian Detection using Adaptive Confidence Thresholding and Inter-Frame Correlation*
Mufleh Al-Shatnawi, Vida Movahedi (York University), Amir Asif (Concordia University), and Aijun An (York University)

Lecture Session LS.T2: Multimedia Quality, Human Factors, and HCI
Session Chair: Giulia Boato (University of Trento)
Time: 4:30pm – 6:00pm

ID Title and Authors

8 *Hybrid-based Facial Expression Recognition Approach for Human-Computer Interaction*
Yacine Yaddaden (Université du Québec à Chicoutimi), Mehdi Adda (Université du Québec à Rimouski), Abdenour Bouzouane, Sebastien Gaboury, and Bruno Bouchard (Université du Québec à Chicoutimi)

18 *Study on Viewing Time with Regards to Quality Factors in Adaptive Bitrate Video Streaming*
Pierre Lebreton, Kimiko Kawashima, Kazuhisa Yamagishi, and Jun Okamoto (NTT Network Technology Laboratories)

59 *User-Independent Detection of Swipe Pressure using a Thermal Camera for Natural Surface Interaction*
Tim Dunn, Sean Banerjee, and Natasha Kholgade Banerjee (Clarkson University)

104 *Video Quality Evaluation for Tile-Based Spatial Adaptation*
Hiba Yousef, Jean Le Feuvre (Telecom ParisTech), Giuseppe Valenzise (Université Paris-Sud), and Vedad Hulusic (Bournemouth University)

115 *Heterogeneous Spatial Quality for Omnidirectional Video*
Hristina Hristova, Xavier Corbillon, Gwendal Simon (IMT Atlantique), Viswanathan Swaminathan (Adobe), and Alisa Devlic (Huawei)

Friday, August 31, 2018

Keynote 3
Time: 9:00am – 10:00am

Title: **Innovations in Assistive Technologies**

Henrique Malvar
Chief Scientist, Microsoft Research

Bio: Henrique (Rico) Malvar is a Microsoft Distinguished Engineer and the Chief Scientist for Microsoft Research. He currently leads a new team at MSR developing technologies for people with disabilities. He joined Microsoft Research in 1997, founding the signal processing group, which developed new technologies such as new media compression formats used in Windows, Xbox, and Office, and microphone array processing technologies used in Windows, Xbox Kinect, and HoloLens. Rico was a key architect for the media compression formats WMA and JPEG XR and made key contributions to the H.264 video format, used by most Web video services. Rico received a PhD from MIT in 1986 and is a Member of the US National Academy of Engineering. He has over 120 issued US patents and over 160 publications. He is an IEEE Fellow and has received many awards, including the Technical Achievement Award from the IEEE Signal Processing Society in 2002.

Abstract: Computing and information technologies have changed our lives. We can't imagine living without our mobile devices, computers, and cloud applications and services. They have increased our productivity and our entertainment options tremendously over the past few decades. Still, for person with disabilities, the full potential of those technologies may be out of reach. There are over one billion people with disabilities in the world, and we have the opportunity and responsibility to leverage technological advances to significantly improve their lives, by making modern technologies more accessible. In this presentation we will discuss some recent advances in assistive technologies that help overcome disabilities in vision, hearing, mobility, and cognition, with examples of some of the work at Microsoft. Those include advances in computer vision, audio signal processing, and natural language processing.

Special Session SS.F1: Multimodal Machine Learning: Advances and Applications
Session Chairs: Qifei Wang (Google)
Time: 10:20am – 12:10pm

ID	Title and Authors
21	*Privacy-Preserving Age Estimation for Content Rating* Linwei Ye (University of Manitoba), Binglin Li (Simon Fraser University), Noman Mohammed (University of Manitoba), Yang Wang (University of Manitoba), and Jie Liang (Simon Fraser University)

24 *Adversarial Attacks on Face Detectors using Neural Net Based Constrained Optimization*
Avishek Bose and Parham Aarabi (University of Toronto)

49 *Fast, Robust, and Accurate Image Denoising via Very Deeply Cascaded Residual Networks*
Lulu Sun, Yongbing Zhang, Xingzheng Wang, Haoqian Wang, and Qionghai Dai (Tsinghua University)

55 *3-Stream Convolutional Networks for Video Action Recognition with Hybrid Motion Field*
Wukui Yang, Shan Gao, Wenran Liu, and Xiangyang Ji (Tsinghua University)

147 *A Robust HER2 Neural Network Classification Algorithm Using Biomarker-Specific Feature Descriptors*
Prerna Singh and Ramakrishnan Mukundan (University of Canterbury)

Poster/Demo Session PS.F1: Multimedia Systems, Tools, and Applications
Session Chair: Yongjun Wu (Amazon)
Time: 1:30pm – 2:30pm

ID Title and Authors

132 *Cloud-Based Tools for Endangered Language Documentation and Analysis*
Archana Dhere and Min Chen (University of Washington, Bothell)

140 *Similar Image Retrieval from X-Ray Database*
Nandinee Fariah Haq (University of British Columbia) and Mehdi Moradi (IBM Almaden Research Center)

160 *Crop Disease Automatic Diagnosis System Based on Smart Mobile Phone and CNN*
Wanjie Liang (Jiangsu Academy of Agricultural Sciences), Zhang Hong (University of Alberta), and Hongxin Cao (Jiangsu Academy of Agricultural Sciences)

161 *Clustering-Based Encoding Adaptation for Video Streaming*
Hai Wei, Yang Yang, Deepthi Nandakumar, Srikanth Kotagiri, Yongjun Wu, Ben Waggoner, Avisar Ten-ami, Bruce Li, and Winston BA (Amazon Video)

162 *A Novel Augmented Reality Framework for Museum Exhibits*
Julien Li-Chee-Ming, Zheng Wu, Randy Tan, Ryan Tan, Naimul Mefraz Khan, Andy Ye, and Ling Guan (Ryerson University)

163 *Automatic Music Accompaniment System Applied to Singing Recreation at Long-Term Geriatric Health-Care Facilities*
Yasuyuki Saito (National Institute of Technology), Yasuji Sakai (Atsugi-city Animated Supporter), Yuu Igarashi (Sound Scape), Eita Nakamura (Kyoto University), Suguru Agata (Asia Pacific Electronic Keyboard Association), and Shigeki Sagayama (Meiji University)

164 *Effects Selection Tool for Improving Visual Attraction of a Target Object*
Natsumi Suzuki and Yohei Nakada (Meiji University)

165 *De-sketching*
Lior Bragilevsky and Ivan V. Bajić (Simon Fraser University)

166 *DFTS: Deep Feature Transmission Simulator*
Harshavardhan Unnibhavi (Indian Institute of Technology (ISM), Dhanbad), Hyomin Choi, Saeed Ranjbar Alvar, and Ivan V. Bajić (Simon Fraser University)

168 *Restricted Live Streaming Genre Identification*
Dong Her Shih, Meng-Yan Lin, and Po-Yuan Shih (National Yunlin University of Science and Technology)

170 A Recurrent Neural Network for Multisensory Non-Intrusive Load Monitoring on a Raspberry Pi
Alon Harell, Stephen Makonin, and Ivan V. Bajić (Simon Fraser University)

Lecture Session LS.F1: Deep Learning for Multimedia Processing
Session Chair: Cha Zhang (Microsoft Research)
Time: 2:50pm – 4:20pm

ID Title and Authors

10 *Rethinking Recurrent Latent Variable Model for Music Composition*
Eunjeong Stella Koh, Shlomo Dubnov, and Dustin Wright (University of California, San Diego)

32 *Deep Siamese Network for Multiple Object Tracking*
Bonan Cuan, Khalid Idrissi, and Christophe Garcia (INSA-Lyon)

56 *Bone Age Assessment with X-ray Images Based on Contourlet Motivated Deep Convolutional Networks*
Xun Chen, Chao Zhang, and Yu Liu (Hefei University of Technology)

88 *A Deep Convolutional Network Based Supervised Coarse-to-Fine Algorithm for Optical Flow Measurement*
Meiyuan Fang, Yanghao Li (Tsinghua University), Yuxing Han (South China Agriculture University), and Jiangtao Wen (Tsinghua University)

108 *Memory-Efficient Deep Salient Object Segmentation Networks on Gridized Superpixels*
Caglar Aytekin, Xingyang Ni, Francesco Cricri, Lixin Fan, and Emre Aksu (Nokia Technologies)